

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :

C12N 15/12; C07K 15/06

A2

(11) International Publication Number:

WO 94/0155

(43) International Publication Date:

20 January 1994 (20.01.94)

(21) International Application Number: PCT/US93/06296

(22) International Filing Date: 1 July 1993 (01.07.93)

(30) Priority data:

07/907,908

2 July 1992 (02.07.92)

US

(71) Applicant: THE SALK INSTITUTE FOR BIOLOGICAL STUDIES [US/US]; 10010 North Torrey Pines Road, La Jolla, CA 92037 (US).

(72) Inventors: EVANS, Ronald, M. ; 8615 La Jolla Scenic Drive, North, La Jolla, CA 92037 (US). MC KEOWN, Michael, B. ; 3083 Fried Avenue, San Diego, CA 92122 (US). ORO, Anthony, E. ; 10710 Escobar Drive, San Diego, CA 92124 (US). SEGRAVES, William, A. ; 416 Meetinghouse Road, Ambler, PA 19002 (US). YAO, Tso-Pang ; 3765-G Miramar Street, La Jolla, CA 92037 (US).

(74) Agent: REITER, Stephen. E.; Pretty, Schroeder, Brueggemann & Clark, 444 South Flower Street, Suite 2000, Los Angeles, CA 90071 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

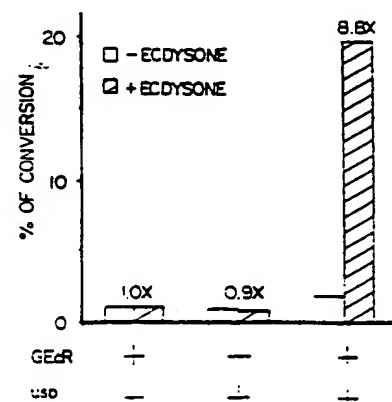
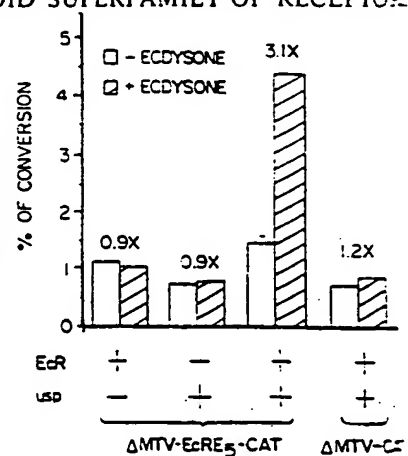
Published

Without international search report and to be republished upon receipt of that report.

(54) Title: MULTIMERIC FORMS OF MEMBERS OF THE STEROID/THYROID SUPERFAMILY OF RECEPTORS WITH THE ULTRASPIRACLE RECEPTOR

(57) Abstract

In accordance with the present invention, it has been discovered that various members of the steroid/thyroid superfamily of receptors can interact with the insect-derived ultraspiracle receptor, to form multimeric species. Accordingly, the interaction of at least one member of the steroid/thyroid superfamily of receptors with at least the dimerization domain of the ultraspiracle receptor modulates the ability of said member of the steroid/thyroid superfamily of receptors to transactivate transcription of genes maintained under hormone expression control in the presence of the cognate ligand for said member of the superfamily.



**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NE	Niger
BE	Belgium	GN	Guinea	NL	Netherlands
BF	Burkina Faso	GR	Greece	NO	Norway
BG	Bulgaria	HU	Hungary	NZ	New Zealand
BJ	Benin	IE	Ireland	PL	Poland
BR	Brazil	IT	Italy	PT	Portugal
BY	Belarus	JP	Japan	RO	Romania
CA	Canada	KP	Democratic People's Republic of Korea	RU	Russian Federation
CF	Central African Republic	KR	Republic of Korea	SD	Sudan
CG	Congo	KZ	Kazakhstan	SE	Sweden
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovak Republic
CM	Cameroon	LU	Luxembourg	SN	Senegal
CN	China	LV	Latvia	TD	Chad
CS	Czechoslovakia	MC	Monaco	TG	Togo
CZ	Czech Republic	MG	Madagascar	UA	Ukraine
DE	Germany	ML	Mali	US	United States of America
DK	Denmark	MN	Mongolia	UZ	Uzbekistan
ES	Spain			VN	Viet Nam
FI	Finland				

MULTIMERIC FORMS OF MEMBERS OF  
THE STEROID/THYROID SUPERFAMILY OF RECEPTORS  
WITH THE ULTRASPIRACLE RECEPTOR

RELATED APPLICATIONS

This application is a continuation-in-part of  
United States Application Serial Number 07/497,935, filed  
5 March 22, 1990, now pending.

FIELD OF THE INVENTION

The present invention relates to interactions  
10 between members of the steroid/thyroid superfamily of  
receptor proteins, novel combinations of various members of  
the steroid/thyroid superfamily of receptor proteins, and  
methods of using such combinations.

15 BACKGROUND OF THE INVENTION

Transcriptional regulation of development and  
homeostasis in complex eukaryotes, including humans and  
other mammals, birds, fish, insects, and the like, is  
20 controlled by a wide variety of regulatory substances,  
including steroid and thyroid hormones. These hormones  
exert potent effects on development and differentiation of  
phylogenetically diverse organisms. The effects of  
hormones are mediated by interaction with specific, high  
25 affinity binding proteins referred to as receptors.

A number of receptor proteins are known, each  
specific for steroid hormones [e.g., estrogens (estr gen  
receptor), progesterones (progesterone receptor),  
30 glucocorticoid (glucocorticoid receptor), androgens  
(androgen receptor), aldosterones (mineralocorticoid  
receptor), vitamin D (vitamin D receptor)], retinoids  
(e.g., retinoic acid receptor) or thyroid hormones (e.g.,  
thyroid hormone receptor). Receptor proteins have been

found to be distributed throughout the cell population of complex eukaryotes in a tissue specific fashion.

Molecular cloning studies have made it possible to demonstrate that receptors for steroid, retinoid and thyroid hormones are all structurally related and comprise a superfamily of regulatory proteins. These regulatory proteins are capable of modulating specific gene expression in response to hormone stimulation by binding directly to cis-acting elements.

It is known that steroid or thyroid hormones, protected forms thereof, or metabolites thereof, enter cells and bind to the corresponding specific receptor protein, initiating an allosteric alteration of the protein. As a result of this alteration, the complex of receptor and hormone (or metabolite thereof) is capable of binding with high affinity to certain specific sites on chromatin. One of the primary effects of steroid and thyroid hormones is an increase in transcription of a subset of genes in specific cell types.

A number of transcriptional control units which are responsive to members of the steroid/thyroid superfamily of receptors have been identified. These include the mouse mammary tumor virus 5'-long terminal repeat (MTV LTR), responsive to glucocorticoid, aldosterone and androgen hormones; the transcriptional control units for mammalian growth hormone genes, responsive to glucocorticoids, estrogens and thyroid hormones; the transcriptional control units for mammalian prolactin genes and progesterone receptor genes, responsive to estrogens; the transcriptional control units for avian ovalbumin genes, responsive to progesterones; mammalian metallothionein gene transcriptional control units, responsive to glucocorticoids; and mammalian hepatic  $\alpha_2$ -globulin gene transcriptional control units, responsive to

androgens, estrogens, thyroid hormones, and glucocorticoids.

A major obstacle to further understanding and more widespread use of the various members of the steroid/thyroid superfamily of hormone receptors has been a lack of awareness of the possible interactions of various members of the steroid/thyroid superfamily of hormone receptors, and an understanding of the implications of such interactions on the ability of members of the steroid/thyroid superfamily of hormone receptors to exert transcriptional regulation of various physiological processes.

DNA binding studies on the glucocorticoid receptor (GR) and the estrogen receptor (ER) have indicated that these receptors bind to their hormone response elements (HREs) as homodimeric complexes [see, for example, Kumar and Chambon in *Cell* 55:145-156 (1988) and Tsi et al., in *Cell* 55:361-369 (1988)]. However, recent biochemical analysis has revealed that some other receptors (including retinoic acid receptor (RAR), thyroid hormone receptor (TR), and the vitamin D receptor (VDR)) can not efficiently bind to cognate response elements as homodimers, but rather require additional factors present in cell nuclear extracts to achieve high affinity DNA binding [see, for example, Murray and Towle in *Mol. Endocrinol.* 3:1434-1442 (1989), Glass et al., in *Cell* 63:729-738 (1990), Liao et al., in *Proc. Natl. Acad. Sci. USA* 87:9751-9755 (1990), and Yang et al., in *Proc. Natl. Acad. Sci. USA* 88:3559-3563 (1991)].

Several recent reports have identified members of the retinoid X receptor family (RXR; see, for example, Mangelsdorf et al., in *Nature* 345:224-229 (1990) and *Gene Dev.* 6:329-344 (1992), and Leid et al., in *Cell* 68:377-395 (1992)) as factors that can interact with RAR and potentiate DNA binding by forming a novel RAR/RXR

heterodimer [see, for example, Yu et al., in Cell 67:1251-1266 (1991), Kliewer et al., in Nature 355:446-449 (1992), Leid et al., supra, and Zhang et al., in Nature 355:441-446 (1992)]. Interestingly, RAR is not the only receptor with which RXR can interact. In fact, RXR has been found to be capable of heterodimerizing with several other members of the nuclear receptor superfamily, including VDR, TR (see Kliewer, et al., supra) and peroxisome proliferator-activated receptor (PPAR; see, for example, Issemann and Green in Nature 347: 645-650 (1990)).

Although the physiological significance of these interactions remains to be definitively determined, the capability of nuclear receptors to heterodimerize suggests the existence of an elaborate network through which distinct nuclear hormone receptor classes are capable of modulating each other's activity. In addition, the possible existence of other factors that can potentially interact with members of the steroid/thyroid superfamily and potentiate DNA binding by forming novel heteromeric species remains to be determined.

#### BRIEF DESCRIPTION OF THE INVENTION

In accordance with the present invention, we have discovered that various members of the steroid/thyroid superfamily of receptors can combine with the insect derived ultraspiracle receptor (or functional fragments thereof comprising at least the dimerization domain thereof) to form a multimeric complex receptor. Accordingly, the combination of a first receptor species with the ultraspiracle receptor (or a truncated form thereof comprising at least the dimerization domain thereof) is capable of modulating the ability of the first receptor species to trans-activate transcription of genes maintained under steroid hormone or hormone-like expression control in the presence of cognate ligand for said first

receptor.

#### BRIEF DESCRIPTION OF THE FIGURES

5           Figure 1 presents a comparison of amino acid identity for various domains of invention ultraspiracle receptor (usp) in comparison with previously identified receptors human RXR-alpha (hRXR $\alpha$ ), human retinoic acid receptor-alpha (hRAR $\alpha$ ) and human glucocorticoid receptor  
10 (hGR).

          Figure 2 presents the % conversion of substrate by chloramphenicol acetyltransferase (CAT) as a result of cotransfection of mammalian (CV1) cells with ecdysone  
15 receptor (EcR) encoding vector and/or ultraspiracle receptor (usp) encoding vector along with CAT reporter vector which contains an ecdysone response element (EcRE).

#### DETAILED DESCRIPTION OF THE INVENTION

20

          In accordance with the present invention, there are provided multimeric receptor species which belong to the steroid/thyroid superfamily of receptors, comprising at least one member of the steroid/thyroid superfamily of  
25 receptors, and the ultraspiracle receptor.

          As employed herein, the term "dimerization domain(s)" of a member of the steroid/thyroid superfamily of receptors refers to that portion (or portions) of the  
30 receptor which is involved in the formation of a given multimeric complex receptor. Dimerization domain(s) typically comprise at least a portion of the carboxy-terminal portion of the receptor (i.e., the carboxy-terminal portion of a receptor with respect to the  
35 DNA-binding domain thereof) and/or at least a portion of the DNA binding domain itself. Multiple domains of a given receptor can act in concert as well as independently.



Combinations contemplated by the present invention can broadly be referred to as "multimeric species", which is intended to embrace all of the various oligomeric forms in which members of the steroid/thyroid superfamily of receptors (including fragments thereof comprising the dimerization domains thereof) are capable of associating with at least the dimerization domain of the ultraspiracle receptor. Thus, reference to "combinations" of steroid receptors or "multimeric" forms of steroid receptors with at least the dimerization domain of the ultraspiracle receptor includes heterodimeric combinations of one member of the steroid/thyroid superfamily of receptors (including fragments thereof comprising the dimerization domain thereof) with at least the dimerization domain of the ultraspiracle receptor, heterotrimeric combinations of one or two members of the steroid/thyroid superfamily of receptors (including fragments thereof comprising the dimerization domains thereof) with at least the dimerization domain of the ultraspiracle receptor, heterotetrameric combinations of one, two or three members of the steroid/thyroid superfamily of receptors (including fragments thereof comprising the dimerization domains thereof) with at least the dimerization domain of the ultraspiracle receptor, and the like.

25

As employed herein, the term "ultraspiracle receptor" refers to a novel invertebrate polypeptide which has a DNA binding domain of about 66 amino acids with at least 9 Cys residues, more than about 75 % amino acid identity in comparison with the DNA binding domain of hRXR-alpha (see Mangelsdorf et al., 1990, supra), less than about 60 % amino acid identity in comparison with the DNA binding domain of hGR, and less than about 60% amino acid identify in comparison with the DNA binding domain of hRAR $\alpha$ . Invention polypeptide can be further characterized by having less than 50% (but typically greater than 40%) amino acid identity in comparison with the ligand binding

domain of hRXR-alpha, but less than 25% amino acid identity in comparison with the ligand binding domains of either hGR or hRAR $\alpha$ . A sequence comparison of amino acid identity between invention receptor and several other receptors is presented in Figure 1.

The deduced amino acid sequence for the ultraspiracle receptor is presented in SEQ ID NO:2 [see also, Oro et al., in Nature 347: 298-301 (1990)]. Also contemplated within the scope of the present invention are peptides comprising a DNA binding domain with substantially the same sequence as that of amino acids 104 - 169 shown in SEQ ID NO:2 (i.e., the DNA binding domain of the ultraspiracle receptor). As employed herein, the term "substantially the same amino acid sequence" refers to amino acid sequences having at least about 80% identity with respect to the reference amino acid sequence, and retaining comparable functional and biological properties characteristic of the protein encoded by the reference amino acid sequence. Preferably, proteins having "substantially the same amino acid sequence" will have at least about 90% amino acid identity with respect to the reference amino acid sequence; with greater than about 95% amino acid sequence identity being especially preferred. Also contemplated within the scope of the present invention are polypeptides having substantially the same sequence as that of amino acids 1 - 513 shown in SEQ ID NO:2. A presently preferred polypeptide of the invention is the polypeptide encoded by vector pXR2C8 [see Oro et al., supra].

As employed herein, the phrase "members of the steroid/thyroid superfamily of receptors" refers to all of the various isoforms of hormone binding proteins that operate as ligand-dependent transcription factors, including members of the steroid/thyroid superfamily of receptors for which specific ligands have not yet been

identified (referred to hereinafter as "orphan receptors").

Each such protein has the intrinsic ability to bind to a specific DNA sequence (i.e., regulatory sequence) associated with the target gene. The transcriptional activity of the gene is modulated by the presence or absence of the cognate hormone (ligand) as a result of binding of ligand to receptor, enabling interaction of receptor with the regulatory sequence.

The DNA-binding domains of all members of this superfamily of receptors are related, consisting of 66-68 amino acid residues, and possessing about 20 invariant amino acid residues, including nine cysteines. A member of the superfamily can be characterized as a protein which contains these diagnostic amino acid residues, which are part of the DNA-binding domain of such known steroid receptors as the human glucocorticoid receptor (amino acids 421-486), the estrogen receptor (amino acids 185-250), the mineralocorticoid receptor (amino acids 603-668), the human retinoic acid receptor (amino acids 88-153), and the like. The highly conserved amino acids of the DNA-binding domain of members of the superfamily are as follows:

```

Cys - X - X - Cys - X - X - Asp* - X -
Ala* - X - Gly* - X - Tyr* - X - X -
X - X - Cys - X - X - Cys - Lys* - X -
Phe - Phe - X - Arg* - X - X - X - X -
X - X - X - X - X - (X - X -) Cys - X -
X - X - X - X - (X - X - X -) Cys - X -
X - X - Lys - X - X - Arg - X - X -
Cys - X - X - Cys - Arg* - X - X -
Lys* - Cys - X - X - X - Gly* - Met
(SEQ ID No 3);

```

wherein X designates non-conserved amino acids within the DNA-binding domain; the amino acid residues noted with an asterisk are residues that are almost universally

conserved, but for which variations have been found in some identified hormone receptors; and the residues enclosed in parenthesis are optional residues (thus, the DNA-binding domain is a minimum of 66 amino acids in length, but can contain several additional residues).

Exemplary members of the steroid/thyroid superfamily of receptors (including the various isoforms thereof) include steroid receptors such as glucocorticoid receptor, mineralocorticoid receptor, progesterone receptor, androgen receptor, vitamin D<sub>3</sub> receptor, and the like; plus retinoid receptors, such as the various isoforms of RAR (e.g., RAR $\alpha$ , RAR $\beta$ , or RAR $\gamma$ ), the various isoforms of RXR (e.g., RXR $\alpha$ , RXR $\beta$ , or RXR $\gamma$ ), and the like; thyroid receptors, such as TR $\alpha$ , TR $\beta$ , and the like; insect derived receptors such as the ecdysone receptor, and the like; as well as other gene products which, by their structure and properties, are considered to be members of the superfamily, as defined hereinabove, including the various isoforms thereof (even though ligands therefor have not yet been identified; such receptors are referred to as "orphan receptors"). Examples of orphan receptors include HNF4 [see, for example, Sladek et al., in *Genes & Development* 4: 2353-2365 (1990)], the COUP family of receptors [see, for example, Miyajima et al., in *Nucleic Acids Research* 16: 11057-11074 (1988), and Wang et al., in *Nature* 340: 163-166 (1989)], COUP-like receptors and COUP homologs, such as those described by Mlodzik et al., in *Cell* 60: 211-224 (1990) and Ladias et al., in *Science* 251: 561-565 (1991), various isoforms of peroxisome proliferator-activated receptors (PPARs; see, for example, Issemann and Green, supra), the insect derived knirps and knirps-related receptors, and the like.

The formation of multimeric receptor(s) can modulate the ability of member(s) of the steroid/thyroid superfamily of receptors to trans-activate transcription of

genes maintained under expression control in the presence of ligand for said receptor. The actual effect on activation of transcription (i.e., enhancement or repression of transcription activity) will vary depending on the receptor species which are part of the multimeric receptor, as well as on the response element with which the multimeric species interacts. Thus, for example, formation of a heterodimer of the ecdysone receptor with the ultraspiracle receptor promotes the ability of the ecdysone receptor to induce trans-activation activity in the presence of an ecdysone response element (see, for example, SEQ ID NO:26).

In accordance with another embodiment of the present invention, there is provided a method to modulate, in an expression system, the transcription activation of a gene by a member of the steroid/thyroid superfamily of receptors in the presence of ligand therefor, wherein the expression of said gene is maintained under the control of a hormone response element, said method comprising:

exposing said system to at least the dimerization domain of the ultraspiracle receptor, in an amount effective to form a multimeric complex receptor with said member of the steroid/thyroid superfamily of receptors.

Exposure of said system to at least the dimerization domain of the ultraspiracle receptor is accomplished by directly administering ultraspiracle receptor (or fragments thereof that allow modification of the receptor through the formation of heterodimeric receptor species) to said system, or by exposing said system to compound(s) and/or condition(s) which induce expression of the ultraspiracle receptor (or dimerization domain thereof). The resulting multimeric receptor species is effective to modulate transcription activation of said gene.

As employed herein, the term "modulate" refers to the ability of a given multimeric complex receptor to either enhance or repress the induction of transcription of a target gene by a given receptor, relative to such ability of said receptor in its uncomplexed state. The actual effect of multimerization on the transcription activity of a receptor will vary depending on the specific receptor species which are part of the multimeric complex receptor, and on the response element with which the multimeric complex receptor interacts. Thus, for example, formation of a heterodimer of the ecdysone receptor and the ultraspiracle receptor provides enhanced trans-activation activity with respect to the ability of the ecdysone receptor alone to promote trans-activation. Conversely, formation of a heterodimer of the ecdysone receptor and the dimerization domain of the ultraspiracle receptor should prevent the ability of ecdysone to promote trans-activation activity, since the resulting multimeric complex receptor will have a reduced ability to bind DNA, relative to the ability of ecdysone-usp multimeric complex to bind DNA.

The term "ecdysone" is employed herein in its generic sense (in accordance with common usage in the art), referring to compounds with the appropriate biological activity, in analogy with the terms estrogen or progestin [see, for example, Cherbas et al., in *Biosynthesis, metabolism and mode of action of invertebrate hormones* (ed. J. Hoffmann and M. Porchet), p. 305-322; Springer-Verlag, Berlin]. 20-Hydroxyecdysone is the major naturally occurring ecdysone. Analogs of the naturally occurring ecdysones are also contemplated within the scope of the present invention, such as for example, ponasterone A, 26-iodoponasterone A, muristerone, inokosterone, 26-mesylinokosterone, and the like.

As employed herein, the phrase "hormone response element" refers to short cis-acting sequences (i.e., having

about 20 bp) that are required for hormonal (or ligand) activation of transcription. The attachment of these elements to an otherwise hormone-nonresponsive gene causes that gene to become hormone responsive. These sequences, commonly referred to as hormone response elements (or HREs), function in a position- and orientation-independent fashion. Unlike other enhancers, the activity of HREs can be modulated by the presence or absence of ligand. See, for example, Evans, Science 240: 889-895 (1988), and the references cited therein. In the present specification and claims, the term "hormone response element" is used in a generic sense to mean and embody the functional characteristics implied by all terms used in the art to describe these sequences.

15

Hormone response elements contemplated for use in the practice of the present invention include naturally occurring response elements as well as modified forms thereof (see, for example, SEQ ID NOs: 7, 12, 15, 25, 26, 28 and 29), as well as synthetic response elements which can be composed of two or more "half sites", wherein each half site comprises the sequence

-RGBNNM-,

wherein

25

R is selected from A or G;

B is selected from G, C, or T;

each N is independently selected from

A, T, C, or G; and

M is selected from A or C;

30

with the proviso that at least 4 nucleotides of said -RGBNNM- sequence are identical with the nucleotides at corresponding positions of the sequence -AGGTCA-, or the half-sites of ecdysone response elements (EcREs) (see, for example, SEQ ID NOs: 26, 28 and 29) and

35

wherein the nucleotide spacing between each of said half-sites falls in the range of 0 up to 15 nucleotides, N.

When one of the half sites described above is incorporated into a synthetic response element in a direct repeat motif, and such half site varies by 2 nucleotides from the preferred sequence of -AGGTCA-, it is preferred that the other half site of the response element be the same as, or vary from the preferred sequence by no more than 1 nucleotide. It is presently preferred that the 3'-half site (or downstream half site) of a pair of half sites vary from the preferred sequence by at most 1 nucleotide.

When the above-described half sites are combined in direct repeat fashion (rather than as palindromic constructs), the resulting synthetic response elements are referred to as "DR-x", wherein "DR" refers to the direct repeat nature of the association between the half sites, and "x" indicates the number of spacer nucleotides between each half site.

Exemplary response elements useful in the practice of the present invention are derived from various combinations of half sites having sequences selected from, for example, -AGGTCA-, -GGTTCA-, -GGGTTA-, -GGGTGA-, -AGGTGA-, -GGGTCA-, and the like.

The nucleotides employed in a non-zero spacer are independently selected from C, T, G, or A.

Exemplary three nucleotide spacers include -AGG-, -ATG-, -ACG-, -CGA-, and the like. Exemplary four nucleotide spacers include -CAGG-, -GGGG-, -TTTC-, and the like. Exemplary five nucleotide spacers include -CCAGG-, -ACAGG-, -CCGAA-, -CTGAC-, -TTGAC-, and the like.



Exemplary response elements contemplated by the present invention include the following DR-3 elements:

5'-AGGTCA-AGG-AGGTCA-3' (SEQ ID No. 4),  
 5'-GGGTGA-ATG-AGGACA-3' (SEQ ID No. 5),  
 5'-GGGTGA-ACG-GGGGCA-3' (SEQ ID No. 6), and  
 5'-GGTTCA-CGA-GGTTCA-3' (SEQ ID No. 7);

the following DR-4 elements:

5'-AGGTCA-CAGG-AGGTCA-3' (SEQ ID No. 8),  
 5'-AGGTGA-CAGG-AGGTCA-3' (SEQ ID No. 9),  
 5'-AGGTGA-CAGG-AGGACA-3' (SEQ ID No. 10),  
 5'-GGGTTA-GGGG-AGGACA-3' (SEQ ID No. 11),

and

5'-GGGTCA-TTTC-AGGTCC-3' (SEQ ID No. 12);

the following DR-5 elements:

5'-AGGTCA-CCAGG-AGGTCA-3' (SEQ ID No. 13),  
 5'-AGGTGA-ACAGG-AGGTCA-3' (SEQ ID No. 14),  
 5'-GGTTCA-CCGAA-AGTTCA-3' (SEQ ID No. 15),  
 5'-GGTTCA-CCGAA-AGTTCA-3' (SEQ ID No. 16),  
 5'-AGGTCA-CTGAC-AGGGCA-3' (SEQ ID No. 17),  
 5'-GGGTCA-TTCAG-AGTTCA-3' (SEQ ID No. 18),

5'-AAGCTTAAG-GGTTCA-CCGAA-AGTTCA-CTCAGCTT-3' (SEQ ID No. 19),

5'-AAGCTTAAG-GGTTCA-CCGAA-AGTTCA-CTCGCATAGCTT-3' (SEQ ID No. 20), and

5'-AAGCTTAAG-GGTTCA-CCGAA-AGTTCA-CTCGCATATATTAGCTT-3' (SEQ ID No. 21);

the ecdysone responsive elements set forth in SEQ ID NOS: 26, 28 and 29, and the like.

Presently preferred response elements contemplated for use in the practice of the present invention include:

5'-AGGTCA-AGG-AGGTCA-3' (SEQ ID No. 4),  
 5'-AGGTCA-CAGG-AGGTCA-3' (SEQ ID No. 8),  
 5'-AGGTGA-CAGG-AGGTCA-3' (SEQ ID No. 9),  
 5'-AGGTCA-CCAGG-AGGTCA-3' (SEQ ID No. 13),  
 5'-AGGTGA-ACAGG-AGGTCA-3' (SEQ ID No. 14),

SEQ ID NOS:26, 28, 29, and the like.

These are especially preferred because they represent synthetic and/or invertebrate sequences which have not been observed in vertebrates, and thus are applicable to a wide variety of reporter systems (i.e., the use of these response elements will not be limited due to any species preference based on the source of the sequence).

In accordance with yet another embodiment of the present invention, there is provided a method to modulate, in an expression system, the transcription activation of a gene by a member of the steroid/thyroid superfamily of receptors in the presence of ligand therefor, and in the further presence of the ultraspiracle receptor, wherein the expression of said gene is maintained under the control of a hormone response element, said method comprising:

exposing said system to compound(s) and/or condition(s) which prevent association of said member with the ultraspiracle receptor or fragments thereof, in an amount effective to prevent said association.

Compound(s) and/or condition(s) which prevent association of said member with the ultraspiracle receptor include hormone-like compounds which act as agonists or antagonists for the ultraspiracle receptor, antibodies raised against the dimerization domain of the ultraspiracle receptor, antibodies raised against the dimerization domain of said member, antisense sequence(s) based on sequence(s) complementary to known RNA encoding at least the dimerization domain of the ultraspiracle receptor, and the like. Amounts of agents effective to prevent such association will vary depending on the particular agents used and can be readily determined by those of skill in the art; typically falling in the sub-nanomolar up to micromolar range.

In accordance with still another embodiment of the present invention, there is provided a method for modulating the expression of an exogenous gene in a subject containing:

5

(i) a DNA construct encoding said exogenous gene under the control of a steroid or steroid-like hormone response element; wherein said response element is not normally present in the cells of said subject,

10

(ii) a receptor which is not normally present in the cells of said subject, wherein said receptor, in the presence of its associated ligand and the ultraspiracle receptor, binds to said steroid or steroid-like hormone response element, and

15

(iii) the ultraspiracle receptor;

20

said method comprising administering to said subject an effective amount of said associated ligand; wherein said ligand is not normally present in the cells of said subject; and wherein said ligand is not toxic to said subject.

25

As employed herein, the term "exogenous" (or "foreign") genes refers to both wild type genes and therapeutic genes, which are introduced into the subject in the form of DNA or RNA, either natural or synthetic. The gene of interest can be introduced into target cells (for in vitro applications), or the gene of interest can be introduced directly into a subject, or indirectly introduced by the transfer of transformed cells into a subject.

30  
35

"Wild type" genes are those that are native to cells of a particular type, but which may be undesirably overexpressed in these cells, or may not be expressed in these cells in biologically significant levels. Thus, for example, while a synthetic or natural gene coding for human insulin would be exogenous genetic material to a yeast cell (since yeast cells do not naturally contain insulin genes), a human insulin gene inserted into a human skin fibroblast cell would be a wild type gene with respect to that cell since human skin fibroblasts contain the genetic material encoding human insulin, although human skin fibroblasts do not express human insulin in biologically significant levels.

Wild type genes contemplated for use in the practice of the present invention include genes which encode a gene product:

the substantial absence of which leads to the occurrence of a non-normal state in said subject; or

a substantial excess of which leads to the occurrence of a non-normal state in said subject; and the like.

As employed herein, the phrase "therapeutic gene" refers to genes which impart a beneficial function to the host cell in which such gene is expressed. Therapeutic genes are those that are not naturally found in host cells. For example, a synthetic or natural gene coding for authentic human insulin would be therapeutic when inserted into a skin fibroblast cell so as to be expressed in a host human, where the host human is not otherwise capable of expressing functionally active human insulin in biologically significant levels.

Therapeutic genes contemplated for use in the practice of the present invention include genes which

encode a gene product:

which is toxic to the cells in which it is expressed; or

5 which imparts a beneficial property to said subject (e.g., disease resistance, etc); and the like.

Exogenous genetic material or genes useful in this embodiment of the present invention include genes that  
10 encode secretory proteins that can be released from said cell; enzymes that can metabolize a substrate from a toxic form to a benign form, or from a benign form to a useful form; regulatory proteins; cell surface receptors; and the like. Such useful genes include, but are not limited to,  
15 genes that encode blood clotting factors such as human factors VIII and IX; genes that encode hormones such as insulin, parathyroid hormone, luteinizing hormone releasing factor (LHRH), alpha and beta seminal inhibins, and human growth hormone; genes that encode proteins such as enzymes,  
20 the absence of which leads to the occurrence of an abnormal state in said subject; genes encoding cytokines or lymphokines such as interferons, granulocytic macrophage colony stimulating factor (GM-CSF), colony stimulating factor-1 (CSF-1), tumor necrosis factor (TNF), and  
25 erythropoietin (EPO); genes encoding inhibitor substances such as alpha<sub>1</sub>-antitrypsin; genes encoding substances that function as drugs, e.g., genes encoding the diphtheria and cholera toxins; and the like.

30 Hormone response elements contemplated for use in this embodiment of the present invention involving modulating the expression of an exogenous gene in a subject include any sequence responsive to the above-described multimeric complex receptors, such as insect response  
35 elements, and the like. See, for example, SEQ ID NOs: 26, 28 and 29.

Insect response elements contemplated for use in modulating the expression of an exogenous gene in a subject according to the present invention include, for example, ecdysone response elements, and the like.

5

Such response elements are operably linked to a suitable promoter for expression of the target gene product. As used herein, the term "promoter" refers to a specific nucleotide sequence recognized by RNA polymerase, the enzyme that initiates RNA synthesis. This sequence is the site at which transcription can be specifically initiated under proper conditions. When exogenous genes, operatively linked to a suitable promoter, are introduced into the cells of a suitable host, the exogenous genes are subject to expression control in the presence of hormone or hormone-like compounds not normally present in the host cells. Exemplary promoters include  $\Delta$ MTV,  $\Delta$ SV,  $\Delta$ ADH promoters, and the like.

20

As employed herein, the phrase "receptor which is not normally present in the cells of said subject" refers to receptors which are not endogenous to the host in which the invention process is being carried out. Receptors which are not endogenous to the host include endogenous receptors modified so as to be non-responsive to ligands which are endogenous to the host in which the invention process is being carried out.

25

Receptor(s) not normally present in the cells of the subject and ultraspiracle receptor (or fragments thereof) can be provided to said subject by direct introduction of the proteins themselves, by introduction of RNA or DNA construct(s) encoding said receptors, by introduction of cells harboring genes encoding said receptor and/or response element, and the like. This can be accomplished in a variety of ways, e.g., by microinjection, retroviral infection, electroporation,

30

35

~~lipofection, and the like.~~

As employed herein, the phrase "associated ligand" refers to a substance or compound which, inside a cell, binds to the receptor protein, thereby creating a ligand/receptor complex, which in turn can bind to an appropriate hormone response element. An associated ligand therefore is a compound which acts to modulate gene transcription for a gene maintained under the control of a hormone response element, and includes compounds such as hormones, growth substances, non-hormone substances that regulate growth, and the like. Ligands include steroid or steroid-like hormones, retinoids, thyroid hormones, pharmaceutically active compounds, and the like. Individual ligands may have the ability to bind to multiple receptors.

In accordance with a still further embodiment of the present invention, there is provided a method of inducing the expression of an exogenous gene in a subject containing:

- (i) a DNA construct encoding an exogenous gene product under the control of a hormone response element; wherein said response element is not normally present in the cells of said subject,
- (ii) DNA encoding a receptor which is not normally present in the cells of said subject, under the control of an inducible promoter; wherein said receptor, in the presence of its associated ligand and the ultraspiracle receptor, binds to said hormone response element,
- (iii) the ultraspiracle receptor, and
- (iv) the associated ligand for said receptor;

said method comprising subjecting a subject to conditions suitable to induce expression of said receptor.

In accordance with yet another embodiment of the present invention, there is provided a method of inducing expression of an exogenous gene product in a subject containing a DNA construct encoding said product under the control of a hormone response element; wherein said response element is not normally present in the cells of said subject, said method comprising introducing into said subject:

a receptor which is not normally present in the cells of said subject; wherein said receptor, in combination with its associated ligand and the ultraspiracle receptor, binds to a hormone response element, activating transcription therefrom,

the ultraspiracle receptor, and the associated ligand for said receptor which is not normally present in the cells of said subject.

In accordance with this embodiment of the present invention, receptor can be provided directly to said subject as the protein, or indirectly by administering to said subject a second DNA construct encoding said receptor, or by administering to said subject cells harboring such constructs. When introduced as part of a second DNA construct, expression of said exogenous gene product and the receptor is preferably maintained under the control of a tissue specific promoter.

In accordance with a further embodiment of the present invention, there is provided a method for the expression of recombinant products detrimental to a host organism, said method comprising:

transforming suitable host cells with:



(i) a construct comprising a sequence encoding said recombinant product under the control of a hormone response element;

5 wherein said response element is not normally present in the cells of said host, and

(ii) DNA encoding a receptor not normally present in said host cells;

10 growing said host cells to the desired level in the substantial absence of hormone(s) which, in combination with said receptor not normally present in the cells of said host and ultraspiracle receptor, is capable of binding to said hormone response element, and

15 inducing expression of said recombinant product by introducing into said host cells the ultraspiracle receptor and hormone(s) which, in combination with said receptor not normally present in the cells of said host, bind to said response element.

20

In one aspect of this embodiment of the invention, wherein host is employed as an expression system for the production of a recombinant product which is toxic to the host, recombinant product is induced only after cell  
25 growth (as opposed to protein expression conditions) has produced a desired density of cell mass. Thus, the desired level of growth in accordance with this embodiment is a level which produces a high cell density, and thereafter expression of product is induced. Conditions suitable for  
30 cell growth (and for protein expression, when desired) can be readily determined by those of skill in the art.

In another aspect of this embodiment of the present invention, wherein the host harbors a DNA construct  
35 as described above, expression of the construct to produce the detrimental product causes ablation of the cells harboring said construct. In this aspect, the desired

level of growth is that level appropriate to ensure the desired distribution of cells harboring the inducible construct. Thus, expression will be induced when it is desired to ablate such cells.

5

As used herein, "ablation" refers to removing or eliminating specific cell types in a culture of a cell population, or in a transgenic animal host by means of a DNA construct that encodes a protein whose presence is not per se toxic to the cells, but which can confer upon the cells a toxic potential due to the ability of the protein to control the expression of substances that are or will become toxic to the cells.

15

The elimination of specific cell-type(s) or specific cell line(s) in accordance with one aspect of the present invention produces a cell population which is substantially free of cells which are not normally present in the wild-type cell population. The elimination of specific cell-type(s) or specific cell line(s), in accordance with another aspect of the present invention, produces a defined altered state in the treated subject.

25

Cell(s) or cell line(s) contemplated to be eliminated in accordance with the present invention can be a cell or cell line capable of providing a desirable component to a cell population, as an exogenous gene product; wherein the ability to eliminate said cell or cell line from said cell population is desired, e.g., once said population achieves the ability to produce sufficient quantities of such component as an endogenous gene product; or, the cell line to be eliminated can be a diseased cell line or a cell line predisposed to a disease state.

35

Normal cell(s) or cell line(s) contemplated to be eliminated in accordance with the present invention are cell(s) or cell line(s), the elimination of which would

result in the creation of a defined altered state in the cell population.

In accordance with a still further embodiment of the present invention, there is provided a method to distinguish the physiological effect of a first hormone receptor in a host from other hormone receptors in said host which respond to the same ligand, said method comprising:

replacing the ligand binding domain of said first receptor with a ligand binding domain from an exogenous receptor to produce a chimeric receptor maintained under the control of a tissue specific promoter;

wherein said exogenous receptor and the ligand to which the exogenous receptor responds are not normally present in said host; and

wherein said exogenous receptor, in the presence of its associated ligand, binds to a hormone response element, thereby activating said response element, and thereafter

monitoring the production of product(s) whose expression is controlled by said first hormone receptor when said host is exposed to ultraspiracle receptor and ligand to which said exogenous receptor responds.

25

In accordance with yet another embodiment of the present invention, there is provided a method to render a mammalian hormone receptor uniquely responsive to a ligand not endogenous to host(s) in which said receptor is normally found, said method comprising:

replacing the ligand binding domain of said receptor with a ligand binding domain from a second receptor;

wherein said second receptor is not normally present in said host; and wherein the ligand to which the second receptor responds is not normally present in said host.

In accordance with a still further embodiment of the present invention, there is provided a method to determine the ligand(s) to which orphan receptor(s) responds, said method comprising:

5 monitoring a host cell containing a reporter construct and a hybrid receptor for expression of product(s) of said reporter construct upon contacting said cell with potential ligands for said orphan receptor and the ultraspiracle receptor;

10 wherein said reporter construct comprises a gene encoding a reporter molecule, operatively linked for transcription to a steroid or steroid-like hormone response element; wherein said response element is not normally present in the cells of

15 said host;

wherein said hybrid receptor comprises:

the N-terminal domain and DNA binding domain of a member of the steroid/thyroid superfamily of

20 receptors, wherein said member is not normally present in the host cells, and wherein said member, in the presence of its associated ligand, binds said response element, activating transcription therefrom, and

25 the ligand binding domain of said orphan receptor.

In accordance with yet another embodiment of the present invention, there is provided an isolated DNA which

30 encodes the ultraspiracle receptor, as described above as well as functional fragments thereof. The complete nucleotide sequence for the ultraspiracle receptor is presented in SEQ ID NO:1 [see also, Oro et al., in Nature 347: 298-301 (1990)]. Also contemplated within the scope

35 of the present invention are sequences encoding polypeptides comprising a DNA binding domain with substantially the same sequence as that of amino acids

104 - 169 shown in SEQ ID NO:2 (i.e., the DNA binding domain of the ultraspiracle receptor). Also contemplated are sequences encoding polypeptides having substantially the same sequence as that of amino acids 1 - 513 shown in  
5 SEQ ID NO:2. Also contemplated are sequences having substantially the same nucleotide sequence as nucleotides 163 - 1704 shown in SEQ ID NO:1. As employed herein, the term "substantially the same as" refers to DNA having at least about 70% homology with respect to the nucleotide  
10 sequence of the DNA fragment with which subject DNA is being compared. Preferably, DNA "substantially the same as" a comparative DNA will be at least about 80% homologous to the comparative nucleotide sequence; with greater than about 90% homology being especially preferred. Also  
15 contemplated are DNAs able to hybridize to the above-described sequences, and having substantially the same functional properties thereof. A presently preferred DNA of the invention is the *EcoRI* fragment of vector pXR2C8 [see Oro, et al., supra].

20

DNA of the invention can optionally be incorporated into expression vector(s) operative in a cell in culture to make the ultraspiracle receptor (or functional fragments thereof) by expression of said DNA in  
25 said cell. For example, the transcription of DNA can be controlled by the *Drosophila melanogaster* actin 5C promoter. Host cells which can be employed for expression of said DNA include *Drosophila melanogaster* Schneider line 2 cells, Kc cells, and the like.

30

The invention will now be described in greater detail by reference to the following non-limiting examples.

EXAMPLESExample I  
Plasmids

5

CMX-EcR was constructed by digesting pActEcR plasmid [Koelle et al., Cell Vol. 67:59-77 (1991)] with HindIII. The resulting HindIII fragment, which contains the EcR coding region, was then inserted into CMXPL1, a derivative of CMX expression vector [Umesono et al. Cell Vol. 65:1255-1266 (1991)]. Expression plasmid CMX-usp was made by inserting the EcoRI fragment from the cDNA clone [Oro et al., Nature Vol. 347:298-301 (1990)] which contains all the usp coding region into CMXPL1 vector.  $\Delta$ MTV-EcRE5-  
15 CAT was constructed by ligation of an EcRE-containing oligonucleotide (SEQ ID NO: 22):

5'-AGCTCGATGG ACAAGTGCAT TGAACCCTTG A  
GCTACC TGTTACGTA ACTTGGGAAC TTCGA

20

into HindIII-cleaved  $\Delta$ MTV-CAT [Hollenberg and Evans, Cell Vol. 55:899-906 (1988)]. Restriction analysis and sequencing of the construct indicated that it contains 5 copies of this oligonucleotide.

25

GEcR was constructed by ligation of a NotI/BamHI fragment containing the DNA and hormone binding domains of a modified EcR cDNA, EcRnx, in place of the DNA and hormone binding domains of the similarly modified glucocorticoid  
30 receptor expression construct pRShGRnx [Giguère et al., Nature Vol. 330:624-629 (1987)]. The modified receptor cDNA was constructed using the site-directed mutagenesis procedure of Kunkel, T.A., Proc. Natl. Acad. Sci. USA Vol. 82:448-492 (1985) to insert NotI (employing SEQ ID NO:23 as  
35 the oligonucleotide template) and XhoI (employing SEQ ID NO:24 as the oligonucleotide template) sites immediately flanking the DNA binding domain. SEQ ID NO:23 is:

5'-CCTGCGCCAC GGCGCCGCC GGAGCTGTG CCTG; and

SEQ ID NO:24 is:

5'-GTGGGTATG CGCCTCGAGT GCGTCGTCCC.

This mutagenesis procedure results in conversion of amino acids 258-260 from ValGlnGlu to ArgProPro and amino acid 5 331 from Pro to Leu.

#### Example II

#### Preparation of receptor protein, cell extracts and gel mobility shift assay

10

To generate protein in vitro, suitable plasmids for human (h)RAR $\alpha$ , hTRB, hVDR, rat PPAR, *Drosophila* (d)usp and dEcR were linearized with restriction enzyme 3' of the termination codon. The linearized templates were used for 15 in vitro transcription and then translation using rabbit reticulocyte lysate according to manufacturer's instruction (Promega). *Drosophila* embryo extract was a gift from Dr. J. Kadonaga and prepared as described by Zoeller et al., in Genes Dev., Vol. 2:68-81 (1988). Schneider cell extracts 20 were prepared following the procedures in Damm et al., Nature Vol. 339:593-597 (1989) and Umesono et al. (1991), supra. The extraction buffer contained 0.4 M KCl in 20 mM HEPES pH 7.5, 20% glycerol, 2 mM DTT and 1 mM PMSF.

25

For gel mobility shift assay, proteins were incubated with binding buffer, which contained 100 mM KCl, 7.5% glycerol, 20 mM HEPES pH 7.5, 2 mM DTT and 0.1% NP-40, on ice for 20 minutes in the presence of 2  $\mu$ g of nonspecific competitor poly dI-dC and other oligo 30 competitors. Then approximately 1 ng of  $^{32}$ P-dCTP probe, which was labelled to specific activity about  $1-5 \times 10^8$  cpm/ $\mu$ g by fill-in reaction with Klenow fragments, was added to the reaction and incubated at room temperature for 20 minutes. Antiserum or preimmune serum was added 10 minutes 35 after the probe was added. The reaction was then loaded into 5% non-denaturing polyacrylamide gel in 0.5 X TBE running buffer [1 X TBE comprises 0.089 M Tris borate,

0.089 M boric acid and 0.002 M EDTA]. After electrophoresis, the gel was dried for autoradiography.

### EXAMPLE III

#### 5                   Preparation of usp antiserum

Primers were designed to amplify the usp coding region which either covered the entire N-terminal and DNA binding domain (from amino acid 1 to 210; GST-uspN) or the  
10 complete coding region (GST-usp) by polymerase chain reaction (PCR). The amplified fragments were subcloned into PGEX2T vector (Pharmacia) for expression in bacteria. The expression of GST fusion protein was performed according to the manufacturer's directions (Pharmacia).

15

The fusion protein GST-uspN was prepared and fractionated on SDS polyacrylamide gel and the band corresponding to the fusion protein was excised. The gel slice was fragmented and used to immunize the rabbit at  
20 three week intervals. The rabbit sera were collected and tested by Western Blot for the ability to recognize usp protein. The positive sera were further purified by the procedure described in Vaughan et al., Met. in Enzymol. (Academic Press Vol. 168:588-617 (1989) with slight  
25 modification. Briefly, full length GST-usp fusion protein and GST protein were purified using Glutathione Sepharose 4B (Pharmacia). The purified proteins were individually coupled to affi-gel 10 according to the manufacturer's protocol (Biorad).

30

To affinity purify the antibody, the crude antiserum was first incubated with GST-coupled affi-gel for 2.5 hours at 4°C with gentle rocking. The unbound fraction was separated from the beads by centrifugation. The  
35 supernatant was then incubated with full length GST-usp coupled affi-gel overnight at 4°C with gentle rocking. The contents were then packed into column and washed with 50 mM



HEPES pH 7.5 supplemented with 0.5 M NaCl. The bound antibodies were eluted by 100 mM glycine. The eluted fractions were neutralized with 1 M Tris pH 8 and pooled, then dialyzed against PBS buffer which contained 0.02% of sodium azide. The purified antiserum was concentrated by Centricon 30 (Amicon) before it was stored. This purified antibody is very specific as it does not cross react with RXRs. It also does not react with other closely related fly nuclear receptors including seven up (svp) type I, II [Mlodzik et al., Cell 60:211-224 (1990)].

#### Example IV

##### Cotransfection assay

Transfection was performed with calcium-phosphate precipitation method as described previously [Umesono et al., Nature 336:262-264 (1989)]. CV1 monkey kidney cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% calf bovine serum. The cells were transfected for 8-9 hours and then the DNA precipitates were washed away and replaced with fresh medium with 10% charcoal-resin double stripped fetal bovine serum. 20-hydroxy-ecdysone (Sigma; 10 mg/ml in ethanol), or ethanol alone was then given to the cells. 24-28 hours later, the cells were harvested. Beta-galactosidase ( $\beta$ Gal) activity was measured and a normalized amount of extract was used for CAT assay (Umesono et al. (1989) supra). The following amount of plasmid DNA was included in the 10 cm plate transfection: 250 ng each of CMX-EcR and RSV-GEcR; 500 ng of CMX-usp; 5  $\mu$ g  $\Delta$ MTV-EcRE<sub>5</sub>-CTA; 5  $\mu$ g of  $\beta$ Gal internal control plasmid CH111 (a derivative of CH110, Promega). The amount of CMX plasmid was kept constant in each transfection by adding CMX-luciferase. PGEM4 was added to bring the total amount of plasmid DNA to 15  $\mu$ g per plate.

EXAMPLE Vusp is the Drosophila nuclear factor that can  
enhance RAR DNA binding activity

5 It has previously been shown that the DNA binding  
activity of bacterially expressed RAR can be enhanced by  
adding cell extracts to the binding reaction [Yang et al.,  
Proc. Natl. Acad. Sci. USA 88:3559-3563 (1991)]. Extracts  
prepared from both mammalian cells and the *Drosophila*  
10 Schneider cell line 2 (S2) had similar effects [Yang et  
al., supra]. The presence of this enhancing activity in  
*Drosophila* cells indicates that a general conserved  
mechanism may be utilized in both mammals and *Drosophila* to  
regulate DNA binding activity of receptors like RAR. To  
15 address this question, experiments were set up to  
characterize this *Drosophila* nuclear activity in S2 cell  
extracts and in embryo extracts.

In gel mobility shift assays using a <sup>32</sup>P-labelled  
20 natural RAR response element - BRARE [see SEQ ID NO:15, see  
also de Thé et al., in Nature Vol. 343:177-180 (1990); and  
Sucov et al., in Proc. Natl. Acad. Sci. USA Vol. 87:5392-  
5398 (1990)] as the probe, in vitro translated RARα was  
incubated with probe under binding conditions either alone  
25 or with 2μg of S2 extracts or embryo extracts. In vitro  
translated RARα, by itself, did not bind with appreciable  
affinity. For mammalian extracts [Glass et al., Cell Vol.  
63:729-738 (1990)], incubating the RARα with either S2 or  
*Drosophila* embryo extracts dramatically enhanced DNA  
30 binding activity, while cell extract alone did not show  
similar binding activity.

Two clues suggested the possibility that the  
observed enhancing activity might be the ultraspiracle  
35 receptor (usp). First, RXR, the putative vertebrate  
homologue of usp, has been shown to enhance RAR DNA binding  
[Yu et al., Cell Vol. 67:1251-1266 (1991); Kliewer et al.,

Nature Vol. 355:446-449 (1992); Leid et al., Cell Vol. 68:377-395 (1992)]. Second, usp protein has been found to be present in both S2 and embryo extracts with relative abundance. Based on these observations, it was  
5 investigated whether usp is the *Drosophila* nuclear activity that can interact with RAR. An affinity purified antibody against usp, prepared as described in Example III, was added to the mobility shift reaction. The affinity purified antibody supershifted the majority of the protein  
10 DNA complex while preimmune serum had no effect in the mobility pattern. By incubating with higher concentration of usp antibody, essentially all the binding complex was supershifted. These results indicate that the activity present in both types of fly extracts can be attributed to  
15 usp protein and that usp is likely the major factor in the extract which interacts with RAR.

To further show that usp is indeed the *Drosophila* component involved in the RAR interactions, usp protein was  
20 in vitro translated in the rabbit reticulocyte lysate (see Example II for procedure employed) and tested whether in vitro translated usp can mimic the fly extracts' activity to interact with RAR $\alpha$ . Neither in vitro translated usp alone, nor RAR $\alpha$  alone, bound to a BRARE probe. However,  
25 when the two proteins were incubated together, a prominent retarded complex appeared. This complex comigrated with the complex detected in the RAR and cell extract mixing experiments. The presence of both usp and RAR protein in this protein/DNA complex was demonstrated by the antibodies  
30 raised against usp and RAR $\alpha$ . Either the affinity purified usp antibody, or the RAR $\alpha$ , specifically affected the protein/DNA complex, while preimmune sera had no effect. This complex likely represents an RAR/usp heterodimer, as its mobility is similar to the RXR/RAR complex which has  
35 been proposed to be a heterodimeric species [Yu et al., supra; Kliewer et al., supra; Leid et al., supra]. Thus it is concluded that usp is the *Drosophila* nuclear activity

which can interact with RAR in binding to specific RARE via the formation of a putative heterodimer.

#### Example VI

#### 5      usp can heterodimerize with several members of the          mammalian nuclear receptor family

The finding that usp is able to heterodimerize with RAR suggested that it would be appropriate to check  
 10 whether this interaction reflected a general ability of usp to form heterodimers with other members of the steroid/thyroid superfamily. Using in vitro translated usp protein, interaction of usp with three mammalian nuclear receptors (TR $\beta$ , VDR and PPAR) was tested in gel mobility  
 15 shift assays with the appropriate response element for each of the three receptors as the probes. Response elements used were as follows:

	<u>Response Element</u>	<u>Abbreviation</u>	<u>Sequence</u>	<u>SEQ IN NO:</u>
20	AOX-PPARE	(DR1) <sup>1</sup>	AGGACA A AGGTCA	25
	SSP1-VDRE	(DR3) <sup>2</sup>	GGTTCA CGA GGTTC	7
	MLV-TRE	(DR4) <sup>3</sup>	GGGTCA TTTC AGGTCC	12
	BRARE	(DR5) <sup>4</sup>	GGTTCA CCGAA AGTTCA	15

- 25 <sup>1</sup>Kliwer et al., submitted to Nature, entitled "9-Cis Retinoic Acid and Peroxisome Proliferator Signalling Pathways Converge Through RXR $\alpha$ -PPAR Interactions"  
<sup>2</sup>Noda et al., Proc. Natl. Acad. Sci. USA 87:9995-9999 (1990)  
<sup>3</sup>Sap et al., Nature 340:242-244 (1989)  
 30 <sup>4</sup>Sucov et al., supra.

For a review providing further discussion with respect to these response elements, see Umesono et al., 1991, supra.

35 With any of the receptors alone (prepared by in vitro translation from cDNA clones) there was very little or no binding to the test probes. However, when they were incubated with usp, a dramatic increase in DNA binding activity could be detected. In a usp dependent fashion,  
 40 TR $\beta$  bound to a natural TR response element (MLV-TRE) and VDR/usp bound to SSP1-VDRE, a natural VDR response element. The usp and PPAR interaction was tested on a PPARE derived

from the acyl CoA oxidase promoter (AOX, Klierer et al., submitted, supra, and references cited therein). PPAR/usp complex bound to AOX-PPARE with high affinity. The usp antibody again showed the presence of usp in those  
5 complexes by shifting the retarded bands in all three combinations, whereas preimmune serum did not affect the binding pattern. Oligonucleotide competition assay demonstrated that the usp dependent heterodimers all showed correct response element specificity. Therefore, by  
10 interacting with usp to form heterodimer, all four mammalian receptors tested achieved high affinity DNA binding to their cognate response elements. It can be concluded, therefore, that the ability of usp to interact with other receptors to form heterodimers is a  
15 characteristic feature of usp, and that receptor heterodimer formation, as exemplified by RXR and usp, is conserved between vertebrates and invertebrates.

#### Example VII

20     Ecdysone receptor and usp can heterodimerize to form a high affinity DNA binding complex

If the substitution of usp for RXR in heterodimerization with RAR and other mammalian receptors  
25 represents a conserved feature of mammalian and *Drosophila* receptors, it can be speculated that there might be one or more *Drosophila* activities which can interact with usp. The ligand binding domain has been shown to contain the dimerization domain for some nuclear receptors [Forman et  
30 al., Mol. Endocrinol. Vol. 3:1610-1626 (1989); Fawell et al., Cell 60:953-962 (1990)] and it is also essential for interaction between RAR and nuclear factors including RXR [Glass et al., supra; Klierer et al. supra]. Sequence comparisons reveal that, with respect to the ligand binding  
35 domain, all RXR heterodimer partners, including RAR, TR and VDR, are much more similar to one another than to other receptors, particularly RXR. Among the *Drosophila* receptor

members, ecdysone receptor (EcR) is one of the receptors which shows strong homology to RAR, VDR and TR within this region. This homology suggests EcR may have an evolutionarily conserved domain that, like RAR, VDR and TR, allows EcR to interact with usp.

To test the potential interaction between usp and EcR, experiments were carried out to determine whether usp is part of the defined EcR DNA binding activity present in ecdysone responsive Schneider 2 cells [Koelle et al. Cell 67:59-77 (1991)]. As shown by Koelle et al., in a gel mobility shift assay using a natural ecdysone response element derived from hsp27 promoter (<sup>32</sup>P-labelled hsp27-EcRE, Riddihough and Pelham, EMBO J. Vol. 6:3729-3734 (1987)) as the probe, one specific major complex could be detected in the S2 extract. This complex can be competed away by specific cold oligonucleotide but not by the unrelated oligonucleotide competitor GREpal, a glucocorticoid response element. To determine whether usp is present in this EcRE binding complex, affinity purified usp antibody was added to the reaction. usp antibody can supershift the specific EcRE binding complex from the S2 extract but not the lower minor complex, which was much less sensitive to the specific cold oligonucleotide competition. Preimmune serum had no effect on the upper major complex but it disrupted the lower minor complex. Antibody against RXR $\alpha$  and RAR $\gamma$  did not affect the specific complex. Therefore, these data demonstrate that usp is part of the EcRE DNA binding complex present in S2 cells, strongly suggesting an interaction between EcR and usp.

The EcR and usp interaction was also tested under more defined conditions. In vitro translated EcR and usp were prepared and their interaction was assayed by gel mobility shift assays. Using the same hsp27-EcRE as the probe, usp did not bind to this element by itself. EcR also failed to bind to hsp27-EcRE, in contrast to the EcRE

binding activity in S2 cells. To test whether usp can complement EcR DNA binding, as it does with mammalian receptors, both usp and EcR were co-incubated in the reaction. In the presence of both receptors, a novel high affinity DNA binding complex appeared. Usp antibody, but not preimmune serum, can supershift the complex, demonstrating that usp is part of the complex. This complex is proposed to be a heterodimeric species, which is consistent with the observation of other usp heterodimers. These data demonstrate that EcR binding to ecdysone response element (hsp27-EcRE) depends upon usp and are consistent with the existence of a functionally significant EcR/usp complex.

15

Example VIIIDNA binding activity of EcR/usp heterodimer is correlated with the ecdysone responsiveness in vivo

To establish that the EcR/usp heterodimer is physiologically relevant, it was set out to determine whether the DNA binding properties of EcR/usp heterodimer can be correlated with ecdysone responsiveness in vivo. This was done by testing EcR/usp heterodimer binding to several wild type and mutant EcREs characterized by their differential ability to mediate the ecdysone responsiveness in cultured cells (for review see Cherbas et al., Genes Dev. Vol. 5:120-131 (1991)). The response elements used in this study are set forth below. The position and the orientation of ERE-like half sites (AGGTCA-like) are marked by arrows. For example, the palindromic motif in hsp27-EcRE is indicated by arrows arranged as → ←. The mutated nucleotides in 11N and 15N-EcRE are in small letters. The arrow in 11N-EcRE covers where the remaining palindromic motif extends. In Eip28/29-EcRE, which is named as dis\*-Eip28/29 in Cherbas et al., supra, the half site which can constitute a highly degenerated palindrome is marked with a wavy, broken line. Note that two ERE half sites are

present in the configuration of direct repeats spaced by three nucleotides in Eip28/29-EcRE. The ability of individual response elements to mediate ecdysone response in cultured cells (summarized from Cherbas et al., supra)  
5 and to serve as high affinity binding site for EcR/usp complex are summarized to the right of the sequences.



38

5

Ability to Mediate Ecdysone Response

Ability to Function as Binding Site for EcR/usp Multimeric Complex

10 hsp27-EcRE (SEQ ID NO:26) ATGGACAAGTGCATTGAACCCCTTGCTCTCT  
 ----->  
 TAACCAGTTCACGTAACTTGGGAACAGAGA  
 <-----

+

15 11N-EcRE (SEQ ID NO:27)

----->  
 atgctGTGCATTGAACgtgctcga  
 tacgaCACGTAACTTgcacgagct  
 <-----

-

20 15N-EcRE (SEQ ID NO:28)

----->  
 atgAAGTGCATTGAACCCgctcga  
 tactTCACGTAACTTGGGcgagct  
 <-----

+

25 Eip28/29-EcRE (SEQ ID NO:29)

----->  
 TAAAGGATCTTGACCCCAATGAACCTTCTTA  
 ATTCCTAGAACTGGGGTTACTTGAAGAAT  
 <----- <-----

+

30

An EcRE derived from the *Drosophila* Eip28/29 gene has been shown to mediate ecdysone response in cultured cells [Cherbas et al., supra]. In contrast to the hsp27-EcRE palindrome, the Eip28/29-EcRE is a composite element containing a direct repeat and a highly degenerated palindromic motif. The ability of EcR/usp complex to recognize this EcRE was examined. This element can effectively compete the EcR/usp binding to hsp27-EcRE. This competition is as effective as that of the hsp27-EcRE itself, demonstrating that the Eip28/29-EcRE is also a high affinity binding site for EcR/usp complex. In contrast, an unrelated competitor (GREpal) had no effect on the DNA binding. The high affinity binding by the EcR/usp complex parallels the ability the Eip28/29-EcRE to mediate ecdysone response in culture cells.

In contrast to the high affinity binding referred to above, a mutant hsp27-EcRE (referred as 11-N-hsp by Cherbas et al., supra, wherein the two nucleotides at the ends of the palindrome and the flanking sequence were changed), failed to serve as a high affinity binding site for the EcR/usp. This oligonucleotide did not compete the specific binding of EcR/usp to hsp27-EcRE, which was consistent with the observation that this mutated EcRE failed to confer ecdysone responsiveness to a heterologous promoter in transfection assay (Cherbas et al., supra). However, a back mutation which regenerated the palindrome motif (15-N-hsp) led to recovery of the ecdysone responsiveness and effective competition for the EcR/usp binding to the wild type hsp27-EcRE. Therefore, the ability of a specific EcRE motif to mediate the ecdysone response in vivo correlated well with its ability to serve as high affinity binding site for EcR/usp complex. These data suggest that the EcR/usp complex can mediate the ecdysone response in vivo.

Example IXusp is present in the embryonic EcRE binding activity

Phenotypic analysis of *usp* reveals that it is  
5 required at a number of developmental events which are  
correlated with known or potential ecdysone-regulated  
processes. For example, *usp* is found to be a necessary  
component for the completion of embryogenesis [Oro et al.,  
"The *Drosophila* retinoid X receptor homolog *ultraspiracle*  
10 functions in both female reproduction and eye  
morphogenesis", Development, in press (1992)]. The  
presence of the ecdysone pulse [Richards, G., Mol. and  
Cell. Endocrinol. Vol. 21:181-197 (1981)] as well as the  
EcR protein (Koelle et al., supra) during embryogenesis  
15 indicates that EcR is also required during embryonic  
development.

Based on the coexpression and functional  
requirement of both activities, it was next tested whether  
20 EcR and *usp* can interact during this developmental stage.  
To this end, the EcRE binding activities in the embryo  
extract were determined by mobility shift assay using 5-10  
 $\mu$ g of embryo extract and  $^{32}$ P-labelled hsp27-EcRE as the  
probe. Specific EcRE binding complexes can be detected, as  
25 demonstrated by specific oligonucleotide and unrelated  
oligonucleotide competition. Indeed, *usp* is present in  
those complexes, as *usp* antibody was able to supershift  
those complexes. Preimmune serum did not alter the  
mobility pattern, although a slightly enhanced overall DNA  
30 binding was observed. The detection of multiple EcRE  
binding complexes is consistent with the existence of  
multiple forms of EcR protein which have been reported  
(Koelle et al., supra). The identities of these complexes  
as EcR complexes were supported by the fact that the upper  
35 complex comigrated with the EcR complex present in the S2  
extract and that the DNA binding specificity of these  
complexes as assayed by oligonucleotide competition was

indistinguishable from the EcR/usp complex prepared in vitro.

Based on these data, it can be concluded that the embryonic EcRE binding complexes contain usp. These data suggest an interaction between endogenous usp and EcR in the *Drosophila* embryo, where both activities are required for embryonic development.

10

#### Example X

#### usp is required for ecdysone responsiveness in heterologous cultured cells

The in vitro DNA binding data suggests that usp is required for EcR high affinity DNA binding. The detection of EcR/usp heterodimeric DNA binding complex from embryo implies that they may interact in vivo. To determine whether usp and EcR can functionally interact in vivo, cotransfection experiments were set up to assay if usp is required for EcR to exert an ecdysone response in cultured cells. Mammalian cells were chosen as the heterologous system because they do not contain endogenous EcR and usp background.

25

Ecdysone response in CV1 cells was determined by using an ecdysone-responsive reporter gene ( $\Delta$ MTV-EcRE<sub>5</sub>-CAT), which contains the same core hsp27-EcRE motif tested in DNA binding assay described earlier. The chloramphenicol acetyl transferase (CAT) activity induced by ecdysone was measured in the presence of different combinations of transfected EcR and usp expression vectors. CAT activity is expressed in Figure 2 as the percentage of conversion which is normalized against the level in EcR alone at the presence of ecdysone assigned as one (column 1, shaded rectangle). The final concentration of 20-hydroxy-ecdysone (Sigma) is 40  $\mu$ M. Columns 1 and 2 show that transfection with EcR alone (column 1) or usp alone (Column 2) does not

respond to 20-OH-ecdysone treatment. Cotransfection of EcR and usp together is shown in column 3 and the induction of CAT activity could be observed by about 3 fold. The requirement of EcRE for induction is demonstrated by using  
5 the parental  $\Delta$ MTV-CAT as reporter. This construct does not respond to ecdysone (column 4).

As shown in Figure 2A, transfection of EcR alone failed to confer responsiveness to 20-hydroxy-ecdysone,  
10 consistent with the idea that EcR by itself cannot mediate ecdysone response. To test whether usp can complement the EcR activity as it did in the DNA binding assay, usp expression vector was cotransfected with EcR into CV1 cells. Cotransfection of usp with EcR indeed was able to  
15 confer a significant response to 20-OH-ecdysone (over 3 fold, Figure 2A). This induction was strictly dependent on both EcR and EcRE, as transfection with usp expression plasmid alone or a reporter construct without EcRE ( $\Delta$ MTV-CAT) did not result in ecdysone responsiveness (Figure 2A).  
20 The level of induction by EcR and usp in CV1 cells is significant but somewhat lower than expected compared with the interaction in DNA binding assay.

To address the possibility that EcR may not  
25 function properly in mammalian cells, the N-terminal domain of the glucocorticoid receptor (GR) was fused to the EcR DNA binding domain and ligand binding domain to create the construct GEcR (see Example 1). The presence of the GR N-terminal domain, which contains a transactivating domain  
30 [Hollenberg and Evans, supra], should increase the transcriptional potency and provide an optimal translation signal for *Drosophila* EcR protein to function properly in the mammalian cells. This construct retained the ability to interact with usp in DNA binding assay. Since usp did  
35 not interact with GR in either DNA binding or in transfection assay, it is clear that GR N-terminal fusion should not affect the basic property of EcR to interact

with usp. A similar GR fusion protein with TR has been shown to behave like wild type TR except that it is a more potent transactivator [Thompson and Evans, Proc. Natl. Acad. Sci. USA Vol. 86:3494-3498 (1989)]. Therefore this  
5 system should increase the sensitivity of a functional interaction between usp and EcR while faithfully retaining the basic property of EcR.

As shown in Figure 2B, expression plasmid of GEcR  
10 was transfected into CV1 cells alone, or with usp expression plasmid. The conversion of CAT was normalized as described for panel A. Note that the scales of the CAT conversion in Figure 2, panels A and B are different. Similar to the results obtained with wild type EcR,  
15 transfection of GEcR alone failed to confer ecdysone response. However, when cotransfected with the usp expression plasmid, the 20-OH-ecdysone treatment induced CAT activity by 8-10 fold (Figure 2B). This induction was also dependent on the presence of the EcRE in the reporter  
20 constructs. Thus, the GEcR mediated ecdysone response is similar to EcR except the signal level is higher (compare Figure 2A and 2B). In conclusion, the presence of usp appears to be essential for EcR and GEcR to exert an ecdysone response in CV1 cells. These data demonstrate  
25 that EcR and usp can interact in vivo and constitute a functional ecdysone response in a heterologous cell line.

While the invention has been described in detail with reference to certain preferred embodiments thereof, it  
30 will be understood that modifications and variations are within the spirit and scope of that which is described and claimed.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: EVANS, RONALD M.  
MCKEOWN, MICHAEL B.  
ORO, ANTHONY E.  
SEGRAVES, WILLIAM A.  
YAO, TSO-PANG

(ii) TITLE OF INVENTION: MULTIMERIC FORMS OF MEMBERS OF THE  
STEROID/THYROID SUPERFAMILY OF RECEPTORS WITH THE  
ULTRASPIRACLE RECEPTOR

(iii) NUMBER OF SEQUENCES: 29

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK  
(B) STREET: 444 South Flower Street, Suite 2000  
(C) CITY: Los Angeles  
(D) STATE: California  
(E) COUNTRY: United States  
(F) ZIP: 90071

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 07/907,908  
(B) FILING DATE: 02-JUL-1992  
(C) CLASSIFICATION:

## (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Reiter,, Stephen E.  
(B) REGISTRATION NUMBER: 31192  
(C) REFERENCE/DOCKET NUMBER: P41 9321

## (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (619) 546-4737  
(B) TELEFAX: (619) 546-9392

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2304 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 163..1701

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGACACGGTG GCGTTGGCAA AGTGAAACCC CAACAGAGAG GCGAAAGCGA GCCAAGACAC  
ACCACATACA CACGAAGAGA ACGAGCAAGA AGAAACCGGT AGCGGGAGGA GCGCGTGCCC

60

120

45

CCAGTTCCTC CAATATACCC AGCACCACAT CACAAGCCCCA GG ATG GAC AAC TGC																174
Met Asp Asn Cys																
1																
GAC	CAG	GAC	GCC	AGC	TTT	CGG	CTG	AGC	CAC	ATC	AAG	GAG	GAG	GTC	AAG	222
Asp	Gln	Asp	Ala	Ser	Phe	Arg	Leu	Ser	His	Ile	Lys	Glu	Glu	Val	Lys	
5					10					15					20	
CCG	GAC	ATC	TCG	CAG	CTG	AAC	GAC	AGC	AAC	AAC	AGC	AGC	TTT	TCG	CCC	272
Pro	Asp	Ile	Ser	Gln	Leu	Asn	Asp	Ser	Asn	Asn	Ser	Ser	Phe	Ser	Pro	
				25					30					35		
AAG	GCC	GAG	AGT	CCC	GTG	CCC	TTC	ATG	CAG	GCC	ATG	TCC	ATG	GTC	CAC	312
Lys	Ala	Glu	Ser	Pro	Val	Pro	Phe	Met	Gln	Ala	Met	Ser	Met	Val	His	
			40					45					50			
GTG	CTG	CCC	GGC	TCC	AAC	TCC	GCC	AGC	TCC	AAC	AAC	AAC	AGC	GCT	GGA	352
Val	Leu	Pro	Gly	Ser	Asn	Ser	Ala	Ser	Ser	Asn	Asn	Asn	Ser	Ala	Gly	
		55					60					65				
GAT	GCC	CAA	ATG	GCG	CAG	GCG	CCC	AAT	TCG	GCT	GGA	GGC	TCT	GCC	GCC	412
Asp	Ala	Gln	Met	Ala	Gln	Ala	Pro	Asn	Ser	Ala	Gly	Gly	Ser	Ala	Ala	
	70					75					80					
GCT	GCA	GTC	CAG	CAG	CAG	TAT	CCG	CCT	AAC	CAT	CCG	CTG	AGC	GGC	AGC	452
Ala	Ala	Val	Gln	Gln	Gln	Tyr	Pro	Pro	Asn	His	Pro	Leu	Ser	Gly	Ser	
85					90					95					100	
AAG	CAC	CTC	TGC	TCT	ATT	TGC	GGG	GAT	CGG	GCC	AGT	GGC	AAG	CAC	TAC	512
Lys	His	Leu	Cys	Ser	Ile	Cys	Gly	Asp	Arg	Ala	Ser	Gly	Lys	His	Tyr	
				105					110					115		
GGC	GTG	TAC	AGC	TGT	GAG	GGC	TGC	AAG	GGC	TTC	TTT	AAA	CGC	ACA	GTG	552
Gly	Val	Tyr	Ser	Cys	Glu	Gly	Cys	Lys	Gly	Phe	Phe	Lys	Arg	Thr	Val	
			120					125					130			
CGC	AAG	GAT	CTC	ACA	TAC	GCT	TGC	AGG	GAG	AAC	CGC	AAC	TGC	ATC	ATA	602
Arg	Lys	Asp	Leu	Thr	Tyr	Ala	Cys	Arg	Glu	Asn	Arg	Asn	Cys	Ile	Ile	
		135					140					145				
GAC	AAG	CGG	CAG	AGG	AAC	CGC	TGC	CAG	TAC	TGC	CGC	TAC	CAG	AAG	TGC	652
Asp	Lys	Arg	Gln	Arg	Asn	Arg	Cys	Gln	Tyr	Cys	Arg	Tyr	Gln	Lys	Cys	
	150					155					160					
CTA	ACC	TGC	GGC	ATG	AAG	CGC	GAA	GCG	GTC	CAG	GAG	GAG	CGT	CAA	CGC	702
Leu	Thr	Cys	Gly	Met	Lys	Arg	Glu	Ala	Val	Gln	Glu	Glu	Arg	Gln	Arg	
165					170					175					180	
GGC	GCC	CGC	AAT	GCG	GCG	GGT	AGG	CTC	AGC	GCC	AGC	GGA	GGC	GGC	AGT	752
Gly	Ala	Arg	Asn	Ala	Ala	Gly	Arg	Leu	Ser	Ala	Ser	Gly	Gly	Gly	Ser	
				185					190					195		
AGC	GGT	CCA	GGT	TCG	GTA	GGC	GGA	TCC	AGC	TCT	CAA	GGC	GGA	GGA	GGA	792
Ser	Gly	Pro	Gly	Ser	Val	Gly	Gly	Ser	Ser	Ser	Gln	Gly	Gly	Gly	Gly	
			200					205					210			
GGA	GGC	GGC	GTT	TCT	GGC	GGA	ATG	GGC	AGC	GGC	AAC	GGT	TCT	GAT	GAC	842
Gly	Gly	Gly	Val	Ser	Gly	Gly	Met	Gly	Ser	Gly	Asn	Gly	Ser	Asp	Asp	
		215					220					225				
TTC	ATG	ACC	AAT	AGC	GTG	TCC	AGG	GAT	TTC	TCG	ATC	GAG	CGC	ATC	ATA	892
Phe	Met	Thr	Asn	Ser	Val	Ser	Arg	Asp	Phe	Ser	Ile	Glu	Arg	Ile	Ile	
	230					235					240					
GAG	GCC	GAG	CAG	CGA	GCG	GAG	ACC	CAA	TGC	GGC	GAT	CGT	GCA	CTG	ACG	942
Glu	Ala	Glu	Gln	Arg	Ala	Glu	Thr	Gln	Cys	Gly	Asp	Arg	Ala	Leu	Thr	
245					250					255					260	



46

TTC CTG CGC GTT GGT CCC TAT TCC ACA GTC CAG CCG GAC TAC AAG GGT	990
Phe-Leu Arg Val Gly Pro Tyr Ser Thr Val Gln Pro Asp Tyr Lys Gly	
265 270 275	
GCC GTG TCG GCC CTG TGC CAA GTG GTC AAC AAA CAG CTC TTC CAG ATG	1038
Ala Val Ser Ala Leu Cys Gln Val Val Asn Lys Gln Leu Phe Gln Met	
280 285 290	
GTC GAA TAC GCG CGC ATG ATG CCG CAC TTT GCC CAG GTG CCG CTG GAC	1086
Val Glu Tyr Ala Arg Met Met Pro His Phe Ala Gln Val Pro Leu Asp	
295 300 305	
GAC CAG GTG ATT CTG CTG AAA GCC GCT TGG ATC GAG CTG CTC ATT GCG	1134
Asp Gln Val Ile Leu Leu Lys Ala Ala Trp Ile Glu Leu Leu Ile Ala	
310 315 320	
AAC GTG GCC TGG TGC AGC ATC GTT TCG CTG GAT GAC GGC GGT GCC GGC	1182
Asn Val Ala Trp Cys Ser Ile Val Ser Leu Asp Asp Gly Gly Ala Gly	
325 330 335 340	
GGC GGG GGC GGT GGA CTA GGC CAC GAT GGC TCC TTT GAG CGA CGA TCA	1230
Gly Gly Gly Gly Gly Leu Gly His Asp Gly Ser Phe Glu Arg Arg Ser	
345 350 355	
CCG GGC CTT CAG CCC CAG CAG CTG TTC CTC AAC CAG AGC TTC TCG TAC	1278
Pro Gly Leu Gln Pro Gln Gln Leu Phe Leu Asn Gln Ser Phe Ser Tyr	
360 365 370	
CAT CGC AAC AGT GCG ATC AAA GCC GGT GTG TCA GCC ATC TTC GAC CGC	1326
His Arg Asn Ser Ala Ile Lys Ala Gly Val Ser Ala Ile Phe Asp Arg	
375 380 385	
ATA TTG TCG GAG CTG AGT GTA AAG ATG AAG CGG CTG AAT CTC GAC CGA	1374
Ile Leu Ser Glu Leu Ser Val Lys Met Lys Arg Leu Asn Leu Asp Arg	
390 395 400	
CGC GAG CTG TCC TGC TTG AAG GCC ATC ATA CTG TAC AAC CCG GAC ATA	1422
Arg Glu Leu Ser Cys Leu Lys Ala Ile Ile Leu Tyr Asn Pro Asp Ile	
405 410 415 420	
CGC GGG ATC AAG AGC CGG GCG GAG ATC GAG ATG TGC CGC GAG AAG GTG	1470
Arg Gly Ile Lys Ser Arg Ala Glu Ile Glu Met Cys Arg Glu Lys Val	
425 430 435	
TAC GCT TGC CTG GAC GAG CAC TGC CGC CTG GAA CAT CCG GGC GAC GAT	1518
Tyr Ala Cys Leu Asp Glu His Cys Arg Leu Glu His Pro Gly Asp Asp	
440 445 450	
GGA CGC TTT GCG CAA CTG CTG CTG CGT CTG CGC CGC TTT GCG ATC GAT	1566
Gly Arg Phe Ala Gln Leu Leu Leu Arg Leu Arg Arg Phe Ala Ile Asp	
455 460 465	
CAG CCT GAA GTG CCA GGA TCA CCT GTT CCT CTT CCG CAT TAC CAG CGA	1614
Gln Pro Glu Val Pro Gly Ser Pro Val Pro Leu Pro His Tyr Gln Arg	
470 475 480	
CCG GCC GCT GGA GGA GCT CTT TCT CGA GCA GCT GGA GGC GCC GCC GCC	1662
Pro Ala Ala Gly Gly Ala Leu Ser Arg Ala Ala Gly Gly Ala Ala Ala	
485 490 495 500	
ACC CGG CCT GGC GAT GAA ACT GGA GTA GGG TCC CGA CTC TAAAGTCGCC	1711
Thr Arg Pro Gly Asp Glu Thr Gly Val Gly Ser Arg Leu	
505 510	
CCCGTTCTCC ATCCGAAAAA TGTTTCATTG TGATTGCGTT TGTTTGCATT TCTCCTCTCT	1771
ATCCCTACAA AAGCCCCCTA ATATTACGCA AAATGTGTAT GTAATTGTTT ATTTTTTTTT	1831

47

TATTACCTAA TATTATTATT ATTATTGATA TAGAAAATGT TTTCCTTAAG ATGAAGATTA	1891
GCCTCCTCGA CGTTTATGTC CCAGTAAACG AAAACAAAC AAAATCCAAA ACTTGAAAAG	1951
AACACAAAAC ACGAACGAGA AAATGCACAC AAGCAAAGTA AAAGTAAAAG TTAAACTAAA	2011
GCTAAACGAG TAAAGATATT AAAATAACGG TTAATAATTAA TGCATAGTTA TGATCTACAG	2071
ACGTATGTAA ACATACAAAT TCAGCATAAA TATATATGTC AGCAGGCGCA TATCTGCGGT	2131
GCTGGCCCCG TTCTAAACCA ATTGTAATTA CTTTTTAACA TAAATTTACC CAAAACGTTA	2191
TCAATTAGAT GCGAGATACA AAAATCACCG ACGAAAACCA ACAAATATA TCTATGTATA	2251
AAAAATATAA GCTGCATAAC AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAA	2304

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Asp	Asn	Cys	Asp	Gln	Asp	Ala	Ser	Phe	Arg	Leu	Ser	His	Ile	Lys	1	5	10	15
Glu	Glu	Val	Lys	Pro	Asp	Ile	Ser	Gln	Leu	Asn	Asp	Ser	Asn	Asn	Ser	20	25	30	
Ser	Phe	Ser	Pro	Lys	Ala	Glu	Ser	Pro	Val	Pro	Phe	Met	Gln	Ala	Met	35	40	45	
Ser	Met	Val	His	Val	Leu	Pro	Gly	Ser	Asn	Ser	Ala	Ser	Ser	Asn	Asn	50	55	60	
Asn	Ser	Ala	Gly	Asp	Ala	Gln	Met	Ala	Gln	Ala	Pro	Asn	Ser	Ala	Gly	65	70	75	80
Gly	Ser	Ala	Ala	Ala	Val	Gln	Gln	Gln	Tyr	Pro	Pro	Asn	His	Pro		85	90		95
Leu	Ser	Gly	Ser	Lys	His	Leu	Cys	Ser	Ile	Cys	Gly	Asp	Arg	Ala	Ser	100	105	110	
Gly	Lys	His	Tyr	Gly	Val	Tyr	Ser	Cys	Glu	Gly	Cys	Lys	Gly	Phe	Phe	115	120	125	
Lys	Arg	Thr	Val	Arg	Lys	Asp	Leu	Thr	Tyr	Ala	Cys	Arg	Glu	Asn	Arg	130	135	140	
Asn	Cys	Ile	Ile	Asp	Lys	Arg	Gln	Arg	Asn	Arg	Cys	Gln	Tyr	Cys	Arg	145	150	155	160
Tyr	Gln	Lys	Cys	Leu	Thr	Cys	Gly	Met	Lys	Arg	Glu	Ala	Val	Gln	Glu	165	170		175
Glu	Arg	Gln	Arg	Gly	Ala	Arg	Asn	Ala	Ala	Gly	Arg	Leu	Ser	Ala	Ser	180	185	190	
Gly	Gly	Gly	Ser	Ser	Gly	Pro	Gly	Ser	Val	Gly	Gly	Ser	Ser	Ser	Gln	195	200	205	

48

Gly	Gly	Gly	Gly	Gly	Gly	Gly	Val	Ser	Gly	Gly	Met	Gly	Ser	Gly	Asn
210						215					220				
Gly	Ser	Asp	Asp	Phe	Met	Thr	Asn	Ser	Val	Ser	Arg	Asp	Phe	Ser	Ile
225					230					235					240
Glu	Arg	Ile	Ile	Glu	Ala	Glu	Gln	Arg	Ala	Glu	Thr	Gln	Cys	Gly	Asp
				245					250					255	
Arg	Ala	Leu	Thr	Phe	Leu	Arg	Val	Gly	Pro	Tyr	Ser	Thr	Val	Gln	Pro
			260					265					270		
Asp	Tyr	Lys	Gly	Ala	Val	Ser	Ala	Leu	Cys	Gln	Val	Val	Asn	Lys	Gln
		275					280					285			
Leu	Phe	Gln	Met	Val	Glu	Tyr	Ala	Arg	Met	Met	Pro	His	Phe	Ala	Gln
	290					295					300				
Val	Pro	Leu	Asp	Asp	Gln	Val	Ile	Leu	Leu	Lys	Ala	Ala	Trp	Ile	Glu
305					310					315					320
Leu	Leu	Ile	Ala	Asn	Val	Ala	Trp	Cys	Ser	Ile	Val	Ser	Leu	Asp	Asp
				325					330					335	
Gly	Gly	Ala	Gly	Gly	Gly	Gly	Gly	Gly	Leu	Gly	His	Asp	Gly	Ser	Phe
			340					345					350		
Glu	Arg	Arg	Ser	Pro	Gly	Leu	Gln	Pro	Gln	Gln	Leu	Phe	Leu	Asn	Gln
		355					360					365			
Ser	Phe	Ser	Tyr	His	Arg	Asn	Ser	Ala	Ile	Lys	Ala	Gly	Val	Ser	Ala
		370				375					380				
Ile	Phe	Asp	Arg	Ile	Leu	Ser	Glu	Leu	Ser	Val	Lys	Met	Lys	Arg	Leu
385					390					395					400
Asn	Leu	Asp	Arg	Arg	Glu	Leu	Ser	Cys	Leu	Lys	Ala	Ile	Ile	Leu	Tyr
				405					410					415	
Asn	Pro	Asp	Ile	Arg	Gly	Ile	Lys	Ser	Arg	Ala	Glu	Ile	Glu	Met	Cys
			420					425					430		
Arg	Glu	Lys	Val	Tyr	Ala	Cys	Leu	Asp	Glu	His	Cys	Arg	Leu	Glu	His
		435					440					445			
Pro	Gly	Asp	Asp	Gly	Arg	Phe	Ala	Gln	Leu	Leu	Leu	Arg	Leu	Arg	Arg
	450					455					460				
Phe	Ala	Ile	Asp	Gln	Pro	Glu	Val	Pro	Gly	Ser	Pro	Val	Pro	Leu	Pro
465					470					475					480
His	Tyr	Gln	Arg	Pro	Ala	Ala	Gly	Gly	Ala	Leu	Ser	Arg	Ala	Ala	Gly
				485					490					495	
Gly	Ala	Ala	Ala	Thr	Arg	Pro	Gly	Asp	Glu	Thr	Gly	Val	Gly	Ser	Arg
			500					505					510		

Leu

49

## (2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 71 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Cys Xaa Xaa Cys Xaa Xaa Asp Xaa Ala Xaa Gly Xaa Tyr Xaa Xaa Xaa  
1 5 10 15  
Xaa Cys Xaa Xaa Cys Lys Xaa Phe Phe Xaa Arg Xaa Xaa Xaa Xaa Xaa  
20 25 30  
Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys  
35 40 45  
Xaa Xaa Xaa Lys Xaa Xaa Arg Xaa Xaa Cys Xaa Xaa Cys Arg Xaa Xaa  
50 55 60  
Lys Cys Xaa Xaa Xaa Gly Met  
65 70

## (2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 15 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AGGTCAAGGA GGTCA

15

## (2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 15 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGGTGAATGA GGACA

15

50

## (2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 15 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGGTGAACGG GGGCA

15

## (2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 15 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGTTCACGAG GTTCA

15

## (2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 16 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AGGTCACAGG AGGTCA

16

## (2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 16 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AGGTGACAGG AGGTCA

16

51

## (2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 16 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AGGTGACAGG AGGACA

## (2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 16 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGGTTAGGGG AGGACA

## (2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 16 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GGGTCATTTC AGGTCC

## (2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 17 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AGGTCACCAG GAGGTCA

52

## (2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AGGTGAACAG GAGGTCA

17

## (2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GGTTCACCGA AAGTTCA

17

## (2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GGTTCACCGA AAGTTCA

17

## (2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AGGTCACTGA CAGGGCA

17

53

## (2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GGGTCATTCA GAGTTCA

17

## (2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 34 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

AAGCTTAAGG GTTCACCGAA AGTTCACTCA GCTT

34

## (2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 38 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

AAGCTTAAGG GTTCACCGAA AGTTCACTCG CATAGCTT

38

## (2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 43 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

AAGCTTAAGG GTTCACCGAA AGTTCACTCG CATATATTAG CTT

43



54

## (2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 62 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: unknown

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

AGCTCGATGG ACAAGTGCAT TGAACCCTTG AGCTACCTGT TCACGTAAC TGGGAAC TTC  
GA

60

52

## (2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 33 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: unknown

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CCTGCGCCAC GCGGCGCCGCGGAGCTGTGC CTG

33

## (2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 29 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: unknown

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GTGGGTATGC GCCTCGAGTG CGTCGTCCC

29

## (2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 13 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: unknown

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

AGGACAAAGG TCA

13

55

## (2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: unknown

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

ATTGGACAAG TGCATTGAAC CCTGTCTCT

30

## (2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: unknown

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

ATGCTGTGCA TTGAACGTGC TCGA

24

## (2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: unknown

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

ATGAAGTGCA TTGAACCCGC TCGA

24

## (2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: unknown

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TAAAGGATCT TGACCCCAAT GAACTTCTTA

30

That which is claimed is:

1. A multimeric receptor comprising at least one member of the steroid/thyroid superfamily of receptors and the ultraspiracle receptor.

5           2. A receptor according to claim 1 wherein said receptor is a heterodimer.

          3. A receptor according to claim 1 wherein said receptor is a heterotrimer.

10

          4. A receptor according to claim 1 wherein said receptor is a heterotetramer.

          5. A receptor according to claim 1 wherein said  
15 member of the steroid/thyroid superfamily of receptors is an insect-derived receptor.

          6. A receptor according to claim 5 wherein said insect-derived receptor is the ecdysone receptor.

20

          7. A receptor according to claim 1 wherein said member of the steroid/thyroid superfamily of receptors is PPAR, VDR, TR $\alpha$ , TR $\beta$ , RAR $\alpha$ , RAR $\beta$  or RAR $\gamma$ .

8. A method to modulate, in an expression system, the transcription activation of a gene by a member of the steroid/thyroid superfamily of receptors in the presence of ligand therefor, wherein the expression of said gene is maintained under the control of a hormone response element, said method comprising:

10 exposing said system to at least the dimerization domain of the ultraspiracle receptor, in an amount effective to form a multimeric complex receptor with said member of the steroid/thyroid superfamily of receptors.

9. A method according to claim 8 wherein the dimerization domain of the ultraspiracle receptor is provided by exposing said system to compound(s) and/or condition(s) which induce expression of a gene encoding said dimerization domain.

10. A method according to claim 8 wherein said member of the steroid/thyroid superfamily of receptors is EcR, PPAR, VDR, TR, or RAR.

11. A method to modulate, in an expression system, the transcription activation of a gene by a member of the steroid/thyroid superfamily of receptors in the presence of ligand therefor, and in the further presence of the ultraspiracle receptor, wherein the expression of said gene is maintained under the control of a hormone response element, said method comprising:

30 exposing said system to compound(s) and/or condition(s) which prevent association of said member with ultraspiracle receptor, or fragments thereof, in an amount effective to prevent said association.

12. A method according to claim 11 wherein said member of the steroid/thyroid superfamily of receptors is selected from EcR, PPAR, RAR, TR, or VDR.

5 13. A method according to claim 11 wherein said compound which prevents association of said member with the ultraspiracle receptor is an anti-ultraspiracle antibody.

10 14. A method for modulating the expression of an exogenous gene in a subject containing:

15 (i) a DNA construct encoding said exogenous gene under the control of a steroid or steroid-like hormone response element; wherein said response element is not normally present in the cells of said subject,

20 (ii) a receptor which is not normally present in the cells of said subject, wherein said receptor, in the presence of its associated ligand and the ultraspiracle receptor, binds to said steroid or steroid-like hormone response element, and

25 (iii) ultraspiracle receptor;

30 said method comprising administering to said subject an effective amount of said associated ligand; wherein said ligand is not normally present in the cells of said subject; and wherein said ligand is not toxic to said subject.

35 15. A method according to Claim 14 wherein said receptor not normally present in the cells of the subject and said ultraspiracle receptor are provided to said subject by DNA construct(s) encoding said receptors.

16. A method according to Claim 15 wherein said receptors are expressed under the control of a tissue specific promoter.

5 17. A method according to Claim 14 wherein said exogenous genes are selected from wild type genes and therapeutic genes.

10 18. A method according to Claim 17 wherein said wild type genes are selected from genes which encode gene products:

the substantial absence of which leads to the occurrence of a non-normal state in said subject; or

15 a substantial excess of which leads to the occurrence of a non-normal state in said subject.

19. A method according to Claim 17 wherein said therapeutic genes are selected from those which encode gene products:

which are toxic to the cells in which they are expressed; or

which impart a beneficial property to said subject.

25

20. A method of inducing the expression of an exogenous gene in a subject containing:

- 5 (i) a DNA construct encoding an exogenous gene product under the control of a hormone response element; wherein said response element is not normally present in the cells of said subject,
- 10 (ii) DNA encoding a receptor which is not normally present in the cells of said subject, under the control of an inducible promoter; wherein said receptor, in the presence of its  
15 associated ligand and the ultraspiracle receptor, binds to said hormone response element,
- (iii) ultraspiracle receptor, and  
20
- (iv) the associated ligand for said receptor which is not normally present in the cells of said subject;
- 25 said method comprising subjecting a subject to conditions suitable to induce expression of said receptor which is not normally present in the cells of said subject.

21. A method of inducing expression of an exogenous gene product in a subject containing a DNA construct encoding said product under the control of a hormone response element; wherein said response element is  
5 not normally present in the cells of said subject, said method comprising introducing into said subject:

a receptor which is not normally present in the cells of said subject; wherein said receptor,  
10 in combination with its associated ligand and ultraspiracle receptor, binds to a hormone response element, activating transcription therefrom,

the ultraspiracle receptor, and  
15 the associated ligand for said receptor.

22. A method for the expression of recombinant products detrimental to a host organism, said method comprising:

transforming suitable host cells with:

20 (i) a construct comprising a sequence encoding said recombinant product under the control of a hormone response element;

wherein said response element is  
25 not normally present in the cells of said host, and

(ii) DNA encoding a receptor not normally present in said host cells;

growing said host cells to the desired level in  
30 the substantial absence of hormone(s) which, in combination with said receptor, is capable of binding to said hormone response element, and

inducing expression of said recombinant product by introducing into said host cells the ultraspiracle  
35 receptor and hormone(s) which, in combination with said receptor not normally present in said host cells, bind to said response element.



23. A method to distinguish the physiological effect of a first hormone receptor in a host from other hormone receptors in said host which respond to the same ligand, said method comprising:

5 replacing the ligand binding domain of said first receptor with a ligand binding domain from an exogenous receptor to produce a chimeric receptor maintained under the control of a tissue specific promoter;

10 wherein said exogenous receptor and the ligand to which the exogenous receptor responds are not normally present in said host; and

15 wherein said exogenous receptor, in the presence of its associated ligand, binds to a hormone response element, thereby activating said response element, and thereafter

20 monitoring the production of product(s) whose expression is controlled by said first hormone receptor when said host is exposed to ultraspiracle receptor and ligand to which said exogenous receptor responds.

24. A method to render mammalian hormone receptor(s) uniquely responsive to a ligand not endogenous to host(s) in which said receptor is normally found, said method comprising:

25 replacing the ligand binding domain of said receptor with a ligand binding domain from a second receptor;

30 wherein said second receptor is not normally present in said host; and wherein the ligand to which the second receptor responds is not normally present in said host.

25. A method to determine the ligand(s) to which orphan receptor(s) responds, said method comprising:

5 monitoring a host cell containing a reporter construct and a hybrid receptor for expression of product(s) of said reporter construct upon contacting said cell with potential ligands for said orphan receptor and the ultraspiracle receptor;

10 wherein said reporter construct comprises a gene encoding a reporter molecule, operatively linked for transcription to a steroid or steroid-like hormone response element; wherein said response element is not normally present in the cells of said host;

15 wherein said hybrid receptor comprises:

20 the N-terminal domain and DNA binding domain of a member of the steroid/thyroid superfamily of receptors, wherein said member is not normally present in the host cells, and wherein said member, in the presence of its associated ligand, binds said response element, activating transcription therefrom, and

the ligand binding domain of said orphan receptor.

25

26. An isolated DNA which encodes a polypeptide, wherein said polypeptide is characterized by having a DNA binding domain of about 66 amino acids with at least 9 Cys residues, wherein said DNA binding domain has:

- 5 (a) more than about 75 % amino acid identity in comparison with the DNA binding domain of hRXR-alpha,
- (b) less than about 60 % amino acid identity in comparison with the DNA binding domain of hGR
- 10 (c) less than about 60% amino acid identity in comparison with the DNA binding domain of hRAR $\alpha$ ,
- as well as functional fragments thereof.

15

27. A DNA according to Claim 26 wherein the polypeptide encoded by said DNA comprises a DNA binding domain with substantially the same sequence as that of amino acids 104-169 shown in SEQ ID NO:2.

20

28. DNA according to Claim 27 wherein the polypeptide encoded by said DNA has substantially the same sequence as that of amino acids 1-513 shown in SEQ ID NO:2.

25

29. DNA according to Claim 28 wherein said DNA comprises a segment with substantially the same nucleotide sequence as nucleotides 163 - 1704 shown in SEQ ID NO:1.

30

30. DNA according to Claim 29 which is pXR2C8.

31. An isolated polypeptide characterized by having a DNA binding domain of about 66 amino acids with at least 9 Cys residues, wherein said DNA binding domain has:

- 5 (a) more than about 75 % amino acid identity in comparison with the DNA binding domain of hRXR-alpha,
- (b) less than about 60 % amino acid identity in comparison with the DNA binding domain of hGR,
- 10 (c) less than about 60% amino acid identity in comparison with the DNA binding domain of hRAR $\alpha$ ,

as well as functional fragments thereof.

15 32. A polypeptide according to Claim 31 wherein said polypeptide comprises a DNA binding domain with substantially the same sequence as that of amino acids 104 - 169 shown in SEQ ID NO:2.

20 33. A polypeptide according to Claim 32 wherein said polypeptide has substantially the same sequence as that of amino acids 1 - 513 shown in SEQ ID NO:2.

25 34. A polypeptide according to Claim 33 wherein said polypeptide is encoded by pXR2C8.

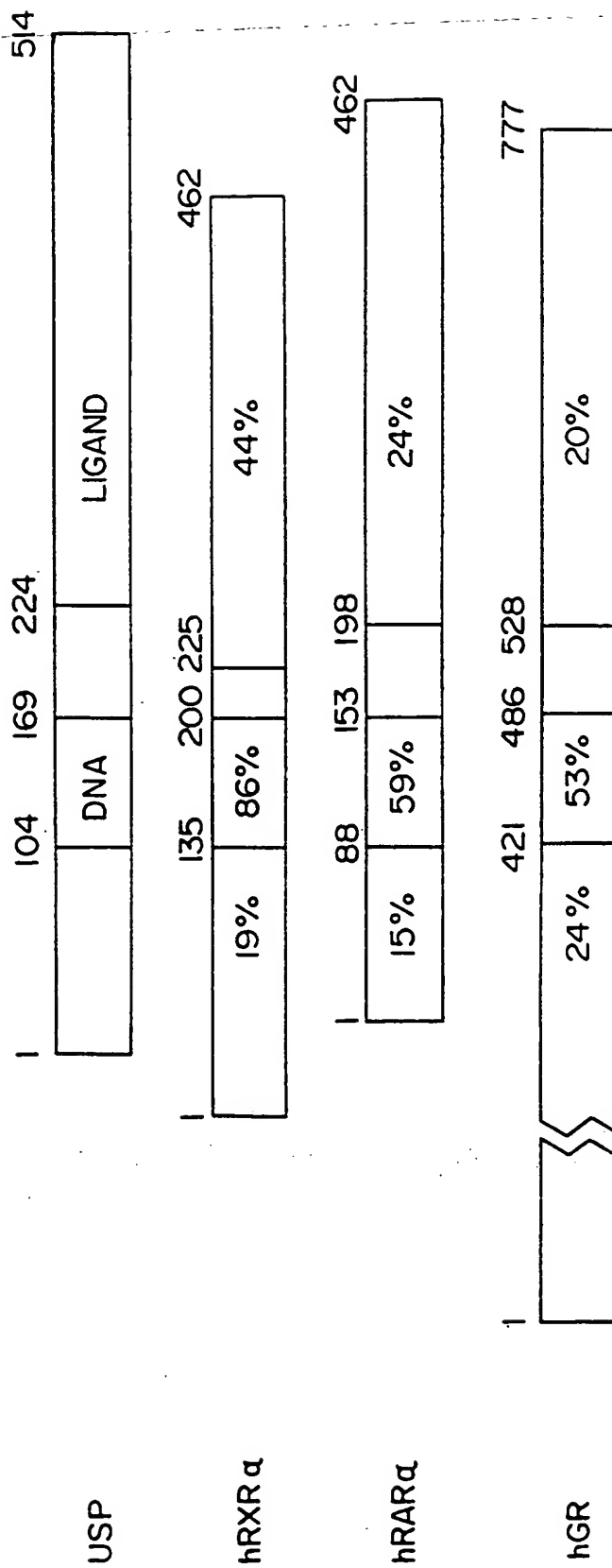


FIG. 1

2 / 2

